

**Table 11. Inter-coder agreement, 1999 study.**

Question	Categories	Kappa between coders 1, 2, and 3			Percent of Inter-coder agreement		
		1 & 2	1 & 3	2 & 3	3 agree	2 agree	0 agree
DNA	all categories	.60	.52	.53	50%	44%	6%
DNA	correct/incorrect	.86	.82	.84	88	12	--
Scientific study	all categories	.52	.34	.42	35	52	13
Scientific study	correct/incorrect	.70	.58	.68	78	22	--
Internet	all categories	.70	.68	.70	75	23	1
Internet	correct/incorrect	.57	.58	.58	85	15	--
Molecule	all categories	.62	.56	.54	57	37	6
Molecule	correct/incorrect	.82	.82	.79	90	10	--
Radiation	all categories	.59	.50	.52	46	42	12
Radiation	correct/incorrect	.73	.66	.66	85	15	--
Experiment	all categories	.81	.77	.78	76	23	1
Experiment	correct/incorrect	.78	.73	.74	82	18	--

## LIMITATIONS

In all survey research there are several possible sources of error, and it is important to recognize these possible sources of error. The primary sources of error connected to the 1999 study are discussed in the following paragraphs.

One source of error in the 1999 study is the exclusion of households without telephones. Approximately 95 percent of households in the United States (excluding group quarters like dormitories, prisons, and hospitals) have a telephone. The presence of a telephone is lowest among **low-income** individuals and families and among non-English speaking groups. Even though the weight procedure makes some correction for this kind of error, it is likely that even within weighting cells, individuals with and without phones may have slightly different attitudinal or demographic profiles.

A second source of error in the 1999 study is the refusal to participate. As was discussed above, approximately a third of possible respondents either directly refused to participate or were able to use an answer machine or other device to avoid talking to an interviewer. Some of this distortion may be corrected by weighting, but it is likely that some of the differences between those who are willing to talk and those who are not is not corrected by weighting.

item comprehension is a third potential source of error in the study. It is likely that some respondents, especially those individuals with little formal education or little exposure to science, may not have fully

understood all of the words and ideas included in the questions. To the maximum extent possible, questions have been constructed to allow respondents to indicate that they did not fully comprehend some items, and interviewers were instructed to denote those interviews where respondents appeared to have a problem with comprehension, but it is likely that some respondents who did not fully comprehend some items still offered a response.

Additionally, some respondents may not have taken the interview seriously, or may have offered insincere responses. The NORC interviewers were trained to identify respondents who seemed to be intoxicated or unable to engage in a conversation and to reschedule those interviews for another time at which the respondent might be more sober or serious. The interviewers also flagged any cases in which the seriousness of the respondent was in doubt. Nevertheless, it is likely that a few respondents offered less than serious responses to some of the items.

A final source of potential error is in the coding of the open-ended items. A number of the questions in the 1999 study involved the collection of open-ended responses from the respondents and the subsequent coding of those responses into numeric codes. All of the coding was closely supervised, as has been described in previous sections, and the inter-coder agreements were high. Additionally, NORC interviewers received extensive training in the capture of the open-ended responses. Nonetheless, it is likely that some of the information included in some of the original responses were not completely entered into the record, or were not coded accurately.

Despite these possible sources of error, the 1999 study reflects the careful use of standard survey research methods and the resulting data provide reasonable estimates of the major attitudinal, behavioral, and knowledge dimensions included in the study.

## **RECOMMENDATIONS**

On the basis of the experience in 1999 and in previous years, we recommend retaining the basic structure of the study. Looking at the time-series record built over the last two decades, a sound time-series has been constructed that is now widely used for a wide range of public policy purposes. Despite increasing response rate problems in most national surveys, we believe that a national study of adult understanding and attitudes still provides essential time-series measures that are relevant to the formulation of public policy. There are, however, some minor changes and improvements that we believe would enhance the value of the study in the future.

### **Sampling**

We are concerned that there is a growing use of cellular telephones and similar instruments as replacements for household telephones, and we recommend that the sample design address this potential source of error. There are conflicting reports in the literature concerning the magnitude of this problem, but it is important that future sample designs explicitly address this problem.